

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:09:02 ; Search time 299.73 Seconds
(without alignments)
26.982 Million cell updates/sec

Title: US-09-331-631a-1_COPY_117_185

Perfect score: 384
Sequence: 1 NR0RDPQOQYEQCKHCORR.....EEQREDEKYEERMKREDN 69

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_15:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	384	100.0	666	09SPL5	09SPL5 macadamia i
2	367	95.6	666	09SPL4	09SPL4 macadamia i
3	363	94.5	625	09SPL3	09SPL3 macadamia i
4	172	44.8	525	10 043358	043358 theobroma c
5	163	42.4	593	10 09SEW4	09SEW4 juglans reg
6	120.5	31.4	554	5 09VPS3	09VPS3 drosophila
7	116.5	30.3	1038	5 060983	060983 dictyosteli
8	115	29.9	1089	12 040947	040947 kaposi's sa
9	114	29.7	810	10 09ZWI3	09ZWI3 cucurbita m
10	107	27.9	1162	12 098148	098148 kaposi's sa
11	105.5	27.5	1129	12 09OR71	09OR71 kaposi's sa
12	105	27.3	1737	5 09TW28	09TW28 dictyosteli
13	104.5	27.2	385	5 017909	017909 caenorhabdi
14	104	27.1	1390	5 077033	077033 dictyosteli
15	103	26.8	838	5 09YOC9	09YOC9 dictyosteli
16	101	26.3	233	4 09UFP2	09UFP2 homo sapien
17	101	26.3	289	4 09NUN6	09NUN6 homo sapien
18	101	26.3	1027	4 09NST7	09NST7 homo sapien
19	101	26.3	1165	4 095819	095819 homo sapien

20	101	26.3	1175	4 075172	075172 homo sapien
21	101	26.3	1233	11 P97820	P97820 mus musculu
22	101	26.3	1257	4 095033	095033 homo sapien
23	100.5	26.2	489	10 09SPL1	09SPL1 glycine max
24	100.5	26.2	2123	5 09U9S7	09U9S7 dictyosteli
25	99.5	25.9	2701	4 09Y520	09Y520 homo sapien
26	98.5	25.7	1268	4 09UKD8	09UKD8 homo sapien
27	98.5	25.7	1276	4 09UKD1	09UKD1 homo sapien
28	98.5	25.7	1297	4 09UKD9	09UKD9 homo sapien
29	98.5	25.7	1305	4 09UKD3	09UKD3 homo sapien
30	98.5	25.7	1323	4 09UKD0	09UKD0 homo sapien
31	98.5	25.7	1331	4 09UKD4	09UKD4 homo sapien
32	98.5	25.7	1333	4 060298	060298 homo sapien
33	98.5	25.7	1352	4 09UKD2	09UKD2 homo sapien
34	98.5	25.7	1360	4 09UKD5	09UKD5 homo sapien
35	98.5	25.7	1700	5 09VVA9	09VVA9 drosophila
36	98	25.5	648	5 09TXB8	09TXB8 dictyosteli
37	98	25.5	876	11 09WU62	09WU62 mus musculu
38	97	25.3	503	5 025777	025777 plasmodium
39	96.5	25.1	1125	5 09W3X0	09W3X0 drosophila
40	96	25.0	479	5 09U380	09U380 caenorhabdi
41	96	25.0	652	4 015410	015410 homo sapien
42	96	25.0	2023	4 075557	075557 homo sapien
43	96	25.0	2023	4 09UND7	09UND7 homo sapien
44	96	25.0	2212	4 09UHV6	09UHV6 homo sapien
45	95.5	24.9	171	11 061118	061118 mus musculu

ALIGNMENTS

```
RESULT 1
09SPL5 PRELIMINARY: PRT: 666 AA.
AC 09SPL5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VICILIN PRECURSOR.
GN AMP2.
OS Macadamia integrifolia (Macadamia nut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OC NCBI_TaxID=60698;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NOT KERNEL.
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antimicrobial peptides is produced by processing of a 7S
RT globulin protein in Macadamia integrifolia kernels.";
DR EMBL: AF161883; AMD54244.1; -.
DR HSSP: P02853; 2PHL.
DR INTERPRO: IPR001113; -.
DR PFM: PF00546; Seedstore_7s; 1.
SQ SEQUENCE 666 AA: 78217 MW; C752B884B2DF0224 CRC64;
```

```
Query Match 100.0%; Score 384; DB 10; Length 666;
Best Local Similarity 100.0%; Pred. No. 5.2e-27;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NR0RDPQOQYEQCKHCORREPRHMOTCCQRCERREYERKKOQRYEEOQREDEKX 60
Db 117 NR0RDPQOQYEQCKHCORREPRHMOTCCQRCERREYERKKOQRYEEOQREDEKX 176
QY 61 EERMKREDN 69
Db 177 EERMKREDN 185
RESULT 2
09SPL4
```

ID	ORGSLD:	PRELIMINARY;	PRT:	666 AA.
DT	01-MAY-2000	(TREMblrel_13, Created)		
DT	01-MAY-2000	(TREMblrel_13, Last sequence update)		
DT	01-OCT-2000	(TREMblrel_13, Last annotation update)		
DE	VICILIN PRECURSOR.			
GN	AMP2.			
OS	Macadamia integrifolia (Macadamia nut).			
OC	Eucaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.			
OX	NCBI_TaxID=60698;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=NOT KERNEL;			
RA	Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;			
RT	"A family of antimicrobial peptides is produced by processing of a 7S			
RT	globulin protein in Macadamia integrifolia.";			
RL	Plant J. 0:0-0(1999).			
DR	EMBL: AF161884; AAD54245.1; "			
DR	HSSD: P02853; 2PBL			
DR	INTERPRO: IPR001113; "			
DR	PFAM: PF00546; Seedstore_7s; 1.			
SO	SEQUENCE	666 AA;	78243 MW;	DECA22E8710F8A7B CRC64;

Query Match	95.68;	Score 367;	DB 10;	Length 666;
Best Local Similarity	95.7%;	Pred. No. 1.7e-25;		
Matches 66;	Conservative 1;	Mismatches 2;	Indels 0;	Gaps 0

```

09 1 NQKRPQOQYECQKNCQRRLEPRNMQCQQRCEPRYEKKRKQOKRYEEOQREDEBK 60
      |||||: || |||||
Db 117 NQKRPQOQYECQKNCQRRLEPRNMQCQQRCEPRYEKKRKQOKRYEEOQREDEBK 176

```

QY	61	ERMKEEDN	69
Db	177	ERMKEEDN	185

RESULT 3
Q9SPL3
ID Q9SPL3
PRELIMINARY;
PRT; 625 AA

```
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VICILIN PRECURSOR (FRAGMENT).
```

OS *Macadamia integrifolia* (Macadamia nut).
 OC Eucaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:
 OC Magnoliophyta: eudicotyledons: Proteaceae: Macadamia.
 OX NCBI_TaxID=60696;

RP SEQUENCE FROM N.A.
RC TISSUE=NOT KERNEL;
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antimicrobial peptides is produced by processing of a 7S

```

RL  Plant J. 0:0-0(1999).
DR  EMBL; AF161885; AAD54246.1; -
DR  HSSP; P02853; 2PHL.
DR  INTERPRO; IPR001113; -.

```

FT	NON_TER	1	1
SQ	SEQUENCE	625 AA;	73586 MW; 415808A89D370296 CRC64;

Query Match	94.58;	Score 363;	DB 10;	Length 625;
Best Local Similarity	-95.78;	Pred. No. 3.7e-25;		
Matches 66; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0

Oy 1 NROKDPQOOYECCOKHCHORRETEPRNMO TC OQRCE RRYEKEKKRKO QRYEEQQ REDEBEKY 60
 | | | | | | | | | | | | | | | | | | |
Db 76 NRGRNDPQOQEYCCKRCOCRREREPNRHNIO IC QRCE RR YE KE KK RK OK QRY EE QQ RE DE BE KY 135

QY	61	EERMKEEDN	69
Db	136	EERMKEGDN	144

RESULT	4	
Q43358		
ID	Q43358	PRELIMINARY; PRU; 525 AA

```
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update,
DE VICILIN PRECURSOR.
```

05 Theobroma cacao (Cacao).
0C Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:
0C Manniophyta: eudicotyledons: core eudicots: Rosidae: eurosids II:
0C Malvales: Malvaceae: Theobroma.

```

RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=LEAVES;
RX      MEDLINE=92288309; PubMed=1600151;

```

RT "Comparison of the structure and nucleotide sequences of vicilin genes
of cocoa and cotton raise questions about vicilin evolution." ;
RL Plant Mol. Biol. 18:1173-1176(1992).
DR EMBL; X62625; CAA44493.1; -

```
DR HSSP; P02853; 2PHL.
DR MENDEL; 30919; Thecc:1188;30919
DR INTERPRO; IPR01113; -.
DR PFAM; PF00546; Seedstore_7s; 1.
```

KM	SIGNAL.		
FT	SIGNAL	1	24
FT	CHAIN	25	525
SQ	SEQUENCE	525 AA;	60798 MW; 19114CD5C248905D CRC64

Query Match	44.8%;	Score 172;	DB 10;	Length 525;
Best Local Similarity	31.4%;	Pred. No. 3.1e-08;		
Matches	32;	Conservative	20;	Mismatches 14;
			Indels	36;
			Gaps	2

[illegible]

Db 95 Q00GQRE00CQKRCWEQYKEQERGEHENYHNHKKNRSEEE 136

09SEW4		
ID	09SEW4	PRELIMINARY;
AC	09SEW4;	PRT;
DT	01-MAY-2000 (TREMblrel, 13, Created)	593 AA

DT 01-OCT-2000 (TrEMBLrel. 15, last annotation update)
DE VILCIN-LIKE PROTEIN PRECURSOR (FRAGMENT).
OS Juglans regia (English walnut).
CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

```

OC  fagales; Juglandaceae; Juglans.
OX  NCBI_TaxID=51240;
RN  [1]
RP  SEQUENCE FROM N.A.

```

RC STRAIN=CV, SUNLAND; TISSUE=SOmatic EMERYO LINE;
RA Teuber S.S., Jarvis K.C., Peterson W.R., Danekar A.M., Ansari A.A.;
RT "Identification and cloning of a cDNA encoding a vicillin-like protein,
Jug 1 2, from English walnut kernel (*Juglans regia*): a major food

RESULT	8
040947	
ID	040947
AC	040947;
DT	01-JAN-1998 (TREMBLrel. 05, Created)
DT	01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE	ORF 73.
OS	Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC	Gammaherpesvirinae; Rhadinovirus.
OX	NCBI_TaxID=37296;
OX	11
RP	SEQUENCE FROM N.A.
RX	MEDLINE=97296220; PubMed=9151804;
RA	Neipel F., Albrecht J.C., Fleckenstein B.;
RT	"Cell-homologous genes in the Kaposi's sarcoma-associated rhadinovirus
RT	human herpesvirus 8: determinants of its pathogenicity?";
RL	J. Virol. 71:4187-4192(1997).
DR	EMBL: U93872; AAB62657.1; -
DR	INTERPRO: IPR002017; -
QO	SEQUENCE 1089 AA; 126232 MW; 48AE2F9837E1A6A CRC64;

Query Match	29.9%	Score 115;	DB 12;	length 1089;
Best Local Similarity	31.3%	Pred. No. 0.0065;		
Matches 21; Conservative	26;	Mismatches 20;	Indels 0;	Gaps 0

[illegible]

RESULT	9			
09ZWM13		PRELIMINARY:	PRT:	810 AA.
ID	09ZWM13			
AC	09ZWM13			
DT	01-MAY-1999	(TREMBLrel. 10, Created)		
DT	01-MAY-1999	(TREMBLrel. 10, Last sequence update)		
DT	01-OCT-2000	(TREMBLrel. 15, Last annotation update)		
DE	PIV100.			
OS	Cucurbita maxima (Pumpkin)	(Winter squash).		
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliopsida; eudicotyledons; core eudicots; Rosidae; eurosids I;			
OC	Cucurbitales; Cucurbitaceae; Cucurbita.			
OX	NCHI_Taxid=3661;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=KURUKAWA AMAKURI NANKIN; TISSUE=COTYLEDON;			
RX	MEDLINE=99107919; PubMed=9891029;			
RA	Maeda K., Shimada T., Kondo M., Nishimura M., Hara-Nishimura I.;			
RT	"Multiple functional proteins are produced by cleaving Asn-Gln bonds			
RT	of a single precursor by vacuolar processing enzyme.";			
RL	J. Biol. Chem. 274:2563-2570(1999).			
DR	EMBL; AB019195; BAA34056.1; "-			
DR	HSSP; P02853; 2PHL.			
DR	INTERPRO; IPR001113; "-			
DR	PFAM; PF00546; Seedstore_7s; 1.			
DR	PRODOM; PD081059; "-1			
QO	SEQUENCE	810 AA; 97314 MW; A829A3F7542266AB CRC64;		

Query Match	29.7%	Score 114;	DB 10;	Length 810;
Best Local Similarly	30.0%	Pred. No. 0.0062;		
Matches 24;	Conservative 14;	Mismatches 30;	Indels 12;	Gaps 1

```
QY      I NKORDPOOYECCQHNCQRRETEPRHMOTCQORCERR-----EKEERKQOKR 48
```

Db 69 NQNSPAEYEVCLRCQVAREGVGEQRKCEHYCEHRLREHQGRGSDVDVEVERADPEME 128
Qy 49 YEEOQREDDEKYEREMKEED 68
 ||||| : | : | : | :
Db 129 REFORRHEHERERRRRERE 148

RESULT	10		
098148			
ID	098148;	PRELIMINARY;	PRT: 1162 AA.
AC	098148;		
DT	01-FEB-1997 (TrEMBLrel. 02, Created)		
DT	01-FEB-1997 (TrEMBLrel. 02, Last sequence update)		
DT	01-JUN-2000 (TrEMBLrel. 14, Last annotation update)		
DE	ORF73 HOMOLOG.		
OS	Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).		
OC	Gammae, dsDNA viruses, no RNA stage; Herpesviridae;		
OC	Gammaherpesvirinae; Rhadinovirus.		
OX	NCBI_TaxID=72296;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
R3	MEDLINE=97048116; PubMed=8892957;		
RA	Ceserman E., Nador R.G., Bal F., Bohanzky R.A., Russo J.J.,		
RA	Moore P.S., Chang Y., Knowles D.M.;		
RT	"Kaposi's sarcoma-associated herpesvirus contains G protein-coupled		
RT	receptor and cyclin D homologs which are expressed in Kaposi's sarcoma		
RT	and malignant lymphoma."		
RL	J. Virol. 70:8218-8223(1996).		

RP SEQUENCE FROM N.A.
RX MEDLINE=97094384; PubMed=8939871;
MOORE P.S., Bashoff C., Weiss R.A., Chang Y.,
RA "Molecular mimicry of human cytokine and cytokine response pathway
RT genes for KSHV";
RL Science 274:1739-1744(1996).
[3]

RP SEQUENCE FROM N.A.
 RP MEDLINE=97121480; PubMed=8962146;
 RA Russo J.J., Bohenzky R.A., Chien M.C., Chen J., Yan M., Meddallena D.,
 RA Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.:
 RT "Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus
 (HHV8)."
 RT Proc. Natl. Acad. Sci. U.S.A. 93:14862-14867(1996).
 RL [4]
 RN
 RN
 RP SEQUENCE FROM N.A.
 RA Russo J.J., Bohenzky R.A., Chien M.C., Chen J., Yan M., Meddallena D.,
 RA Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.:
 RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.
 RL [3]
 RN
 RN
 RP SEQUENCE FROM N.A.
 RA Russo J.J., Bohenzky R.A., Chien M.C., Chien J., Yan M., Meddallena D.,
 RA Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.:
 RL Submitted (May-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U55064; AAC55944.1; -;
 DR EMBL: U75698; AAC57158.1; -;
 DR INTERPRO: IPR002017; -;
 QO SEQUENCE 1162 AA; 135213 MW; 1A72CE01C1CB081C CRC64;

Query Match	27.9%	Score 107;	DB 12;	Length 1162;
Best Local Similarity	31.3%	Pred. No. 0.035;		
Matches	21;	Conservative	25;	Mismatches 21; Indels 0; Gaps 0

QY 2 RQRDPQQYYEQCCCKHCQRRRETEPRHMYTCQQRCEBRYEKEKRRKQDQRYEEQDQREDEBYE 61
::: ||| | : : : | :: :||: |:|: : : :
Db 680 QQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDD 739

```
QY      62  ERMKEED 68
        1 : : : :
Db      740  EQQQQDE 746
```

RESULT 1.1

